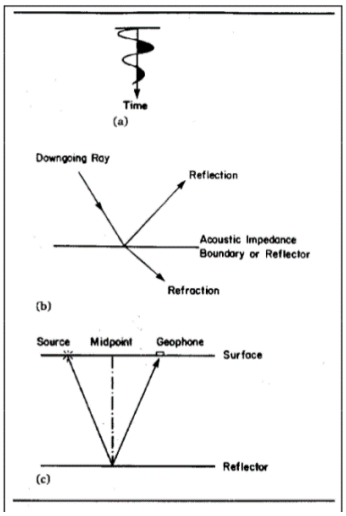
BAB II
TEORI DASAR

2.1 Konsep Dasar Seismik Refleksi

Seismik refleksi merupakan salah satu metode geofisika yang digunakan dalam eksplorasi khususnya hidrokarbon dengan memanfaatkan gelombang elastis sebagai media interpretasi. Pada prinsipnya, metode refleksi memanfaatkan gelombang elastis yang diinjeksikan oleh source yang kemudian merambat melalui suatu medium (bumi) dan hasil dari rambatan tersebut berupa respon yang ditangkap oleh *receiver* seperti pada Gambar 2.1. Respon tersebut akan memberikan informasi seputar litologi bawah permukaan. Informasi yang dibawa tersebut diterjemahkan menjadi bentuk waktu rambat (*travel time*) yang akan memberikan gambaran tentang kecepatan rambat gelombang pada lapisan batuan yang dilalui, amplitudo, frekuensi dan variasi fasa.



Gambar 2. 1 Unsur dasar metode seismik refleksi. (a) Skema wavelet sumber. (b) Refleksi dan refraksi pada batas IA. (c) Geometri refleksi pada reflector horizontal (Sukmono,1999)

Sinyal yang diterima *receiver* dalam satu seismik trace terdiri dari beberapa *wiggle*. Trace seismik merupakan data seismik yang terekam oleh satu *receiver*. Setiap trace merupakan konvolusi sederhana dari reflektivitas bumi dengan fungsi sumber seismik ditambah dengan *noise* (Russel,1996).

$$S(t) = W(t) * r(t) + n(t) \quad (2.1)$$

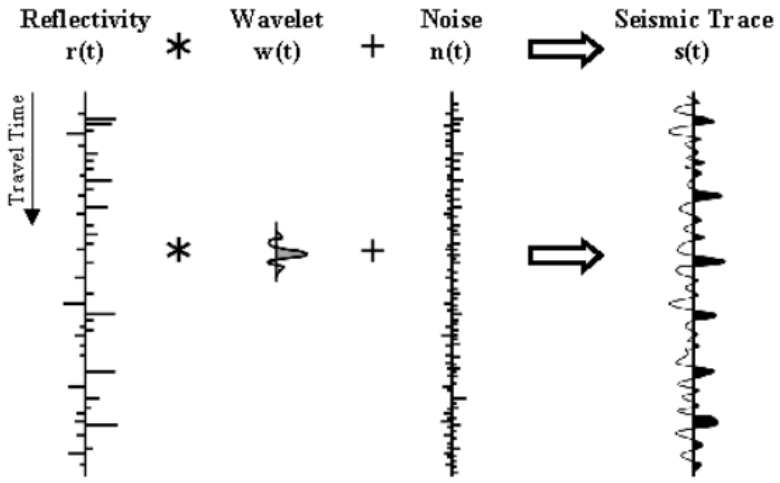
dimana: S(t) = trace seismik

W(t) = wavelet seismik r

(t) = reflektivitas bumi

n(t) = *noise*

* = simbol dari operasi konvolusi



Gambar 2. 2 Ilustrasi proses hasil trace seismik (Sukmono,2001)

2.2 Komponen Seismik Refleksi

2.2.1 Impedansi Akustik

Impedansi akustik adalah fungsi perkalian antara densitas media (ρ) dengan kecepatan media gelombang (V) yang dinyatakan dalam rumus:

$$AI = \rho . V \quad (2.2)$$

Kecepatan memiliki peran yang lebih penting dalam mengontrol nilai AI karena perubahan kecepatan lebih signifikan daripada perubahan densitas secara lateral maupun vertikal. Perubahan nilai impedansi akustik dapat menandakan perubahan karakteristik batuan seperti litologi, porositas, kekerasan, dan kandungan fluida. AI

dapat dianalogikan berbanding lurus terhadap kekerasan batuan dan berbanding terbalik dengan porositas

2.2.2 Koefisien Refleksi

Koefisien refleksi adalah suatu nilai yang mempresentasikan bidang batas antara dua medium yang memiliki impedansi akustik yang berbeda. Untuk gelombang yang mengenai batas lapisan pada umumnya, koefisien refleksinya dapat ditulis :

$$KR = \frac{(AI_2 - AI_1)}{(AI_1 + AI_2)} \quad (2.3)$$

Dimana:

KR = Koefisien refleksi

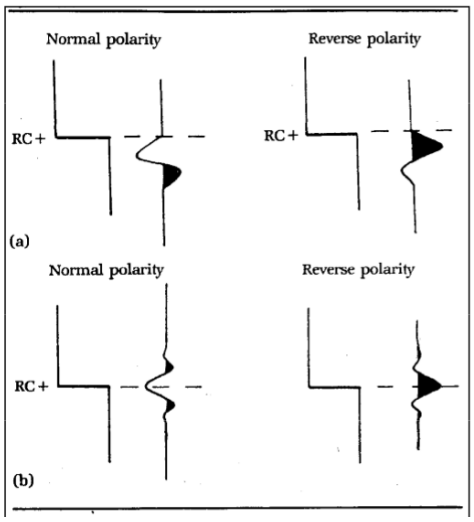
AI_1= Impedansi akustik lapisan atas

AI_2= Impedansi akustik lapisan bawah

2.2.3 Polaritas

Pada polaritas wavelet ini menggambarkan suatu koefisien refleksi positif atau negatif hal ini disebabkan adanya ketidakpastian pada bentuk gelombang seismik yang terekam sehingga perlu dibuat pendekatan suatu polaritas yang berbeda diantaranya polaritas normal dan terbalik *(reverse)*. Meskipun penggunaan kata polaritas hanya mengacu pada perekaman, tapi dalam rekaman seismik penentuan polaritas sangat penting. Dalam penggunaan konvensi *Society of Exploration Geophysicists* (SEG) pada polaritas normal ini, dalam sebuah penampang seismik dengan tampilan polaritas normal kita akan mengharapkan :

1. Bidang batas refleksi dimana $AI_2 > AI_1$ akan berupa *peak.*
2. Bidang batas refleksi dimana $AI_2 < AI_1$ akan berupa *trough.*



Gambar 2. 3 Contoh konvensi polaritas menurut SEG. (a) Fasa minimum (b) zero-phase (Sukmono,2001)

2.2.4 Resolusi Vertikal Seismik

Resolusi adalah jarak minimum antara dua objek yang dapat dipisahkan oleh gelombang seismik (Sukmono, 1999). Resolusi dalam gelombang seismik didefinisikan sebagai kemampuan gelombang seismik untuk memisahkan dua objek yang berbeda. Resolusi vertikal dapat dihitung dalam persamaan

$$rv = \frac{1}{4}\lambda = \frac{v}{4f} \qquad (2.4)$$

Keterangan :

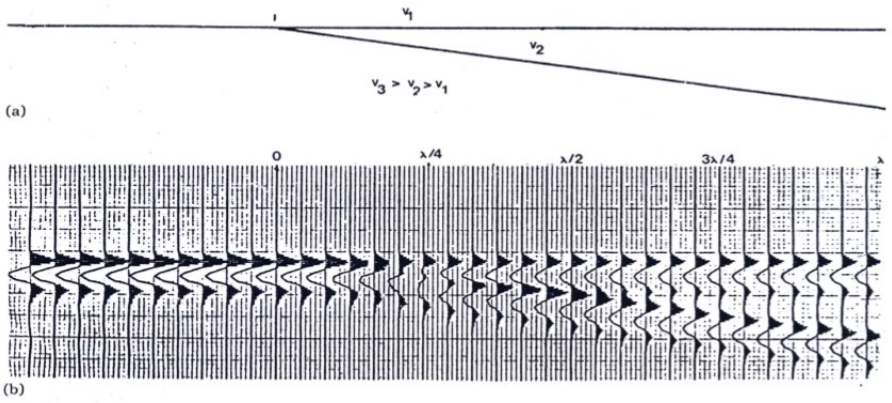
rv = Resolusi vertikal (m)

λ = Panjang gelombang (m)

v = Kecepatan rata-rata (m/s)

f = Frekuensi (Hz)

Ketebalan minimum suatu objek bervariasi untuk dapat memberikan refleksi sendiri antara 1/8 λ sampai 1/30 λ . Resolusi tubuh batuan setara dengan 1/4 λ dalam waktu bolak-balik atau TWT. Hanya batuan yang memiliki ketebalan diatas 1/4 λ yang dapat dibedakan gelombang seismik. Gambar 2.4 menunjukkan bagaimana refleksi ketebalan dua litologi yang berbeda.

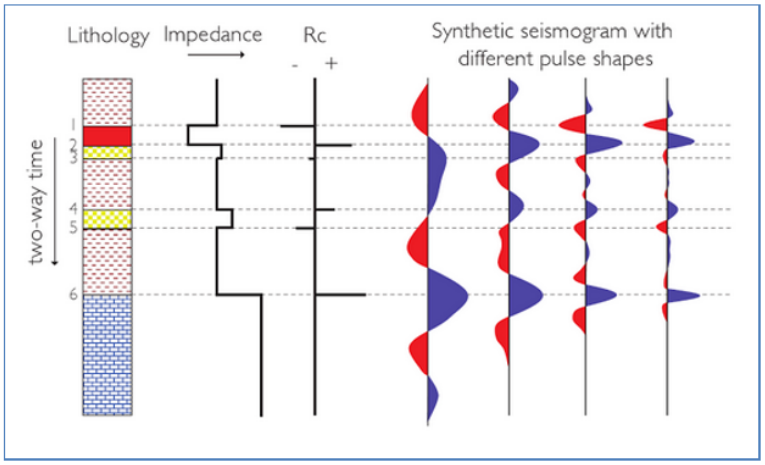


Gambar 2. 4 Refleksi dari tubuh batuan yang mempunyai AI diantara batuan diatas dan dibawahnya. (a) Model. (b) Penampang seismik (Sheriff,1975)

2.3 Pengikatan data seismik dan sumur

2.3.1 Seismogram Sintetik

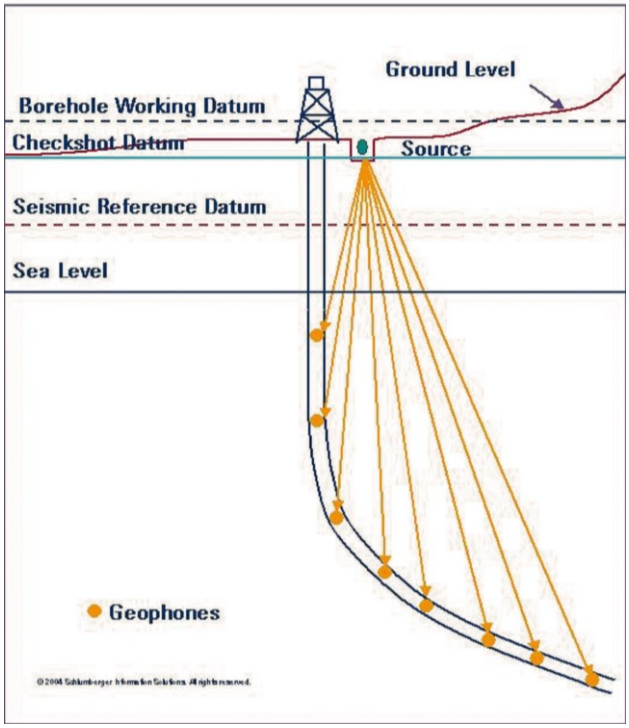
Seismogram sintetik adalah rekaman seismik buatan yang dibuat berdasarkan data log kecepatan dan densitas. Data kecepatan dan densitas membentuk fungsi koefisien refleksi yang kemudian dikonvolusikan dengan wavelet. Seismogram sintetik (Gambar 2.5) dibuat untuk mengkorelasikan antara informasi sumur (litologi, umur, kedalaman, dan sifat-sifat fisis lainnya) terhadap trace seismik guna memperoleh informasi yang lebih lengkap dan akurat. Proses pengikatan seismogram sintetik (*well seismic tie)* dilakukan dengan mengkorelasi data sumur dengan data seismik dengan membutuhkan bantuan data *checkshot*. Pengikatan seismogram sintetik dilakukan karena adanya perbedaan domain, data seismik dalam domain waktu sedangkan data sumur dalam domain kedalaman.



Gambar 2. 5 Seismogram Sintetik yang diperoleh dari konvolusi KR dan wavelet (Rob and Mike, 2014)

2.3.2 Checkshot

Selain itu terdapat data lain yang berguna dalam pengikatan data sumur yaitu data *checkshot*. Data ini berguna untuk mengkonversi data sumur dari domain kedalaman menjadi domain waktu. Apabila data asli checkshot hanya terdapat pada satu sumur saja maka korelasi yang dilakukan yaitu dengan cara memilih event-event target pada sintetik seismogram dan mengesernya pada posisi event-event data seismik atau disebut sebagai shifting. Dalam hal ini digunakan *wavelet* yang diekstrak secara *ricker* dengan menggunakan minimum *phase wavelet*. Hal ini dilakukan dengan cara menemukan posisi serta window target yang akan diekstrak yang kemudian akan dilakukan *shifting*. Hasil *well seismic tie* tergolong baik apabila hasil dari *time shift* mendekati 0 atau sama dengan 0 dan hasil korelasi mendekati 1.



Gambar 2. 6 Survei Checkshot (Brown, 2005)

Survei ini memiliki kesamaan dengan akuisisi data seismik pada umumnya namun posisi *geophone* diletakkan sepanjang sumur bor, atau dikenal dengan survei *Vertical Seismik Profilling (VSP)*. Sehingga data yang didapatkan berupa *one way time* yang dicatat pada kedalaman yang ditentukan sehingga didapatkan hubungan antara waktu jalar gelombang seismik pada lubang bor tersebut.

2.4 Well Logging

2.4.1 Gamma Ray Log

Gamma ray log merupakan salah satu jeni log *spontaneous measurements.* Gamma ray log mengukur radioaktif alami dari formasi sehingga dapat digunakan untuk mengidentifikasi litologi, melakukan korelasi, serta menghitung volume shale. Batu pasir yang tidak memiliki kandungan shale memiliki kandungan material radiokatif yang sedikit sehingga memberikan nilai gamma ray yang rendah. Seiring dengan kandungan shale yang meningkat, maka nilai gamma ray log juga akan semakin meningkat. Hal ini disebabkan oleh adanya peningkatan konsentrasi material radioaktif dalam shale. Walau demikian batu pasir yang bersih juga dapat

memberikan bacaan gamma ray log yang tinggi ketika batu pasir tersebut mengandung potassium feldspar, mika, glaukoit atau uranium.

2.4.2 Sonic log

Sonic log merupakan salah satu dari pengukuran log yang dapat juga digunakan untuk mengukur porositas. Log ini mengukur interval waktu tempuh yang dibutuhkan oleh gelombang kompresional untuk melalui suatu formasi sepanjang sumbu lubang bor. Peralatan sonic log terdiri dari satu atau lebih transmitter gelombang ultra sonic, dan beberapa receiver. Sonic log dipengaruhi oleh dua hal, yaitu litologi dan porositas dari batuan, dimana keduanya merupakan fungsi dari modulus bulk(K), modulus shear serta densitas.

2.4.3 Log Bulk density

Prinsip kerja log density yaitu suatu sumber radioaktif dari alat pengukur di pancarkan sinar gamma dengan intensitas energi tertentu menembus formasi/batuan. Batuan terbentuk dari butiran mineral, mineral tersusun dari atom-atom yang terdiri dari proton dan elektron. Partikel sinar gamma bertumbukan dengan elektron- elektron dalam batuan. Ada dua jenis nilai densitas yang berbeda yang digunakan dalam log densitas, yaitu bulk density (ρ_b atau RHOB) dan matriks densitas (ρ_{ma}). Bulk densitas merupakan densitas dari seluruh formasi, baik bagian padat maupun begian fluida. Sedangkan matriks densitas merupakan densitas dari kerangka solid batuan. Log densitas ini dapat digunakan untuk mengidentifikasi mineral evaporit, mendeteksi kehadiran gas zone, menentukan densitas hidrokarbon, serta mengevaluasi *shaly-sand* reservoir dan litologi kompleks.

2.4.4 Log Resistivity (RT)

Resistivity adalah kemampuan batuan untuk melewatkan arus listrik yang mengalir padanya. Log Resistivity digunakan untuk mendeterminasi zona Mining Area dan zona air, mengindikasikan zona permeabel dengan mendeterminasi porositas resistivitas. Karena batuan dan matrik tidak konduktif, maka kemampuan batuan

untuk menghantarkan arus listrik tergantung pada fluida dan pori. Untuk tujuan geologi, log resistivitas yang digunakan harus diketahui kemampuan resolusinya. Laterolog mampu memberikan gambaran lapisan pada skala yang tepat untuk indikasi batas lapisan, tetapi penggunaannya harus digunakan dan dikorelasikan dengan log lainnya. Log induksi memberikan resolusi batas lapisan yang sangat buruk, tetapi pada saat yang sama semua efek lapisan dirata- rata sedemikian rupa untuk membuat tren litologi menonjol.

2.4.5 Log Neutron

Log neutron merekam Hidrogen index (HI) dari formasi. HI merupakan indikator kelimpahan kandungan hidrogen dalam formasi (dengan asumsi atom H berasal dari HC atau Air). Satuan pengukuran dinyatakan dalam satuan PU (Porosity Unit) (Rider, 1996). Tanggapan alat neutron terutama untuk mencerminkan banyaknya atom hidrogen di dalam formasi. Semakin banyak atom H dalam formasi, maka partikel neutron yang kembali semakin sedikit. Besarnya porositas batuan sama dengan jumlah energi neutron yang hilang, karena atom hidrogen berkonsentrasi pada pori yang terisi fluida (air atau minyak). Pori yang terisi oleh gas akan memiliki pola kurva log neutron akan lebih rendah dari yang seharusnya (*gas effect*). Hal ini terjadi karena konsentrasi hidrogen dalam gas lebih kecil dibandingkan pada minyak dan air (Harsono, 1997).

2.4.6 Interpretasi estimasi Porositas

Menurut Harsono (1997), porositas didefinisikan sebagai rasio antara ruang pori pada batuan dengan volume total batuan. Biasanya diekspresikan dalam satuan persen (%). Pada formasi renggang (*unconsolidated formation*) besarnya porositas tergantung pada distribusi ukuran butiran, tidak pada ukuran butiran mutlak. Kemudian pada porositas yang lebih rendah partikel-partikel batuan umumnya bergabung bersama material yang mengandung silika atau zat kapur, menghasilkan formasi rapat (*consolidated formation*) dengan porositas mendekati nol.

Ada beberapa alat untuk menentukan porositas menurut Riyan (2012) antara lain,

a. Density Log

Dalam menentukan porositas batuan dipengaruhi juga oleh litologi kandungan fluida batuan. Porositas dari log densitas biasanya dinotasikan dengan ΦD yang mempunyai harga sesuai dengan persamaan dibawah ini:

$$\Phi_D = \frac{\rho_{ma} - \rho_{bulk}}{\rho_{ma} - \rho_f} \quad (2.6)$$

Dimana

Φ_D = Porositas Densitas

ρ_{ma} = Densitas matriks batuan (gr/cc), batupasir 2,65

ρ_b = Densitas bulk yang dibaca pada kurva log RHOB (gr/cc)

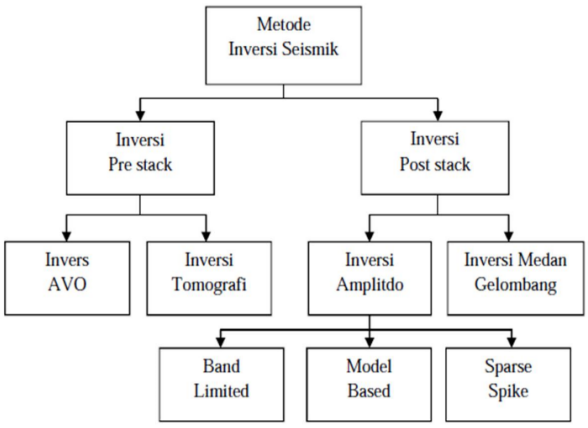
ρ_f = Densitas fluida (air) dibaca dari log header (gr/cc)

Tabel 2. 1 Tabel Nilai porositas berdasarkan kualitas secara umum. (Koesoemadinata dalam Nurwidyanto dkk., 2005)

Nilai Porositas	Kualitas (umum)
0% - 5%	Diabaikan *(negligible)*
5% - 10%	Buruk *(poor)*
10% - 15%	Cukup *(fair)*
15% - 20%	Baik *(good)*
20% - 25%	Sangat Baik *(very good)*
>25%	Istimewa *(excellent)*

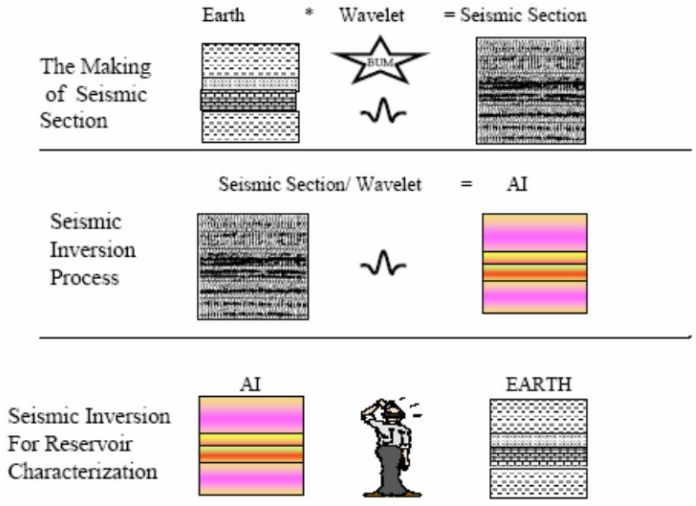
2.4 Metode Inversi Seismik

Secara umum inversi seismik dapat dipisahkan menjadi dua jenis yaitu inversi *pre-stack* dan inversi *post-stack* (gambar 2.7). Inversi seismik *post-stack* merupakan teknik untuk mendapatkan kembali nilai koefisien refleksi dari rekaman seismik yang selanjutnya digunakan untuk menentukan nilai impedansi akustik lapisan batuan. Inversi *post-stack* inilah yang dilakukan pada penelitian ini.



Gambar 2. 7 Berbagai macam metode inversi seismik (Sukmono, 2001)

Proses inversi melakukan pemodelan impedansi akustik dengan cara mengkonvolusi nilai wavelet dengan seismogram sintetis yang sesuai dengan trace penampang seismik. Dimana nilai sintetik seismogram didapat dari data sumur. Hasil dari seismik inversi berupa model impedansi akustik, yang merupakan perubahan secara lateral nilai kecepatan dan massa jenis yang berada disekitar sumur. Model impedansi akustik merupakan salah satu sebuah atribut seismik yang berguna untuk mengkarakterisasi reservoir.



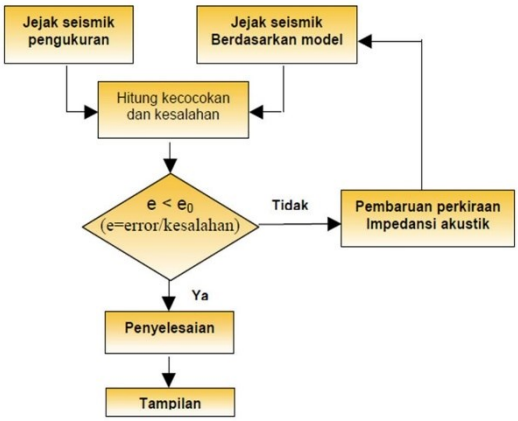
Gambar 2. 8 Ilustrasi Proses Seismik Inversi (Sukmono, 2000)

Inversi seismik adalah suatu proses untuk menghitung model impedansi bawah permukaan yang sesuai dengan penampang seismik. Dua tipe data yang diperlukan

dalam input proses inversi adalah data seismik dan data model inisial yang dibuat pada tahap pembuatan model. Model ini menggambarkan model inisial dari struktur kecepatan yang akan digunakan untuk membatasi inversi. Hasil akhir adalah profil impedansi yang berubah sekecil mungkin dari model inisial namun juga mirip dalam memodelkan data sebenarnya.

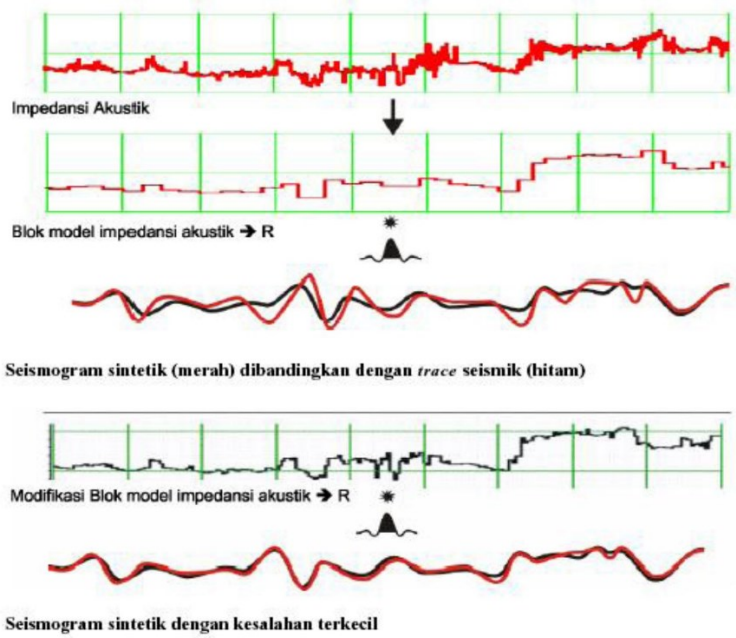
a. Inversi Berbasis Model (*Model Based Inversion*)

Di antara ketiga jenis metode inversi amplitudo, metode inversi *model based* dengan menggunakan teknik inversi *Generalized Linear Inversion* (GLI) memiliki hasil dengan ralat yang terkecil. Menurut Russell (1991), proses untuk menghasilkan model impedansi akustik yang paling cocok dengan data hasil pengukuran berdasarkan nilai rata-rata kesalahan terkecil (*least square*) merupakan proses inversi linear umum (GLI).



Gambar 2. 9 Diagram Metode Inversi Model Based (Russel,2008)

Konsep inversi pada metode ini dimulai dengan membuat model inisial impedansi akustik dengan ukuran blok yang telah ditentukan. Koefisien refleksi diturunkan dari impedansi akustik kemudian dikonvolusikan dengan wavelet yang menghasilkan seismogram sintetik pada tiap-tiap *trace*. Seismogram sintetik ini selanjutnya dibandingkan dengan trace seismik sebenarnya lalu dihitung kesalahannya. Impedansi akustik hasil modifikasi model awal inilah yang merupakan hasil akhir inversi. Proses ini dilakukan secara *iteratif* dengan memodifikasi *blok trace* model hingga diperoleh hasil sintetik dengan kesalahan terkecil.



Gambar 2. 10 Proses inversi berbasiskan model (Russel,2008)

Metode inversi *model based* mempunyai keunggulan karena hasil yang didapatkan memiliki kontrol yang baik karena menghindari inversi langsung dari data seismik. Hasil inversi digambarkan dalam bentuk *z* yang memiliki nilai impedansi akustik yang kontras, sehingga mempermudah dalam penentuan batas suatu lapisan reservoir. Kelemahan inversi *model based* terletak pada ketidakunikan inversi. Dengan kata lain, ada banyak kemungkinan solusi model untuk dapat menghasilkan suatu keluaran hasil yang sama.

2.5. Seismik Multi Atribut

Analisis multiatribut seismik merupakan terminologi yang mencakup semua metode geostatistik yang menggunakan lebih dari satu atribut seismik untuk memprediksi properti fisik bumi (Russel et al.,1997). Metode multi-atribut geostatistik ini dapat dibagi menjadi tiga sub-kategori utama yaitu:

1. Suatu metode perluasan dari co-krigging untuk melibatkan lebih dari satu atribut sekunder untuk memprediksi parameter utama.
2. Metode yang menggunakan matriks kovariansi untuk memprediksi suatu parameter dan atribut input yang telah diberi bobot secara linear.

3. Metode yang menggunakan Artificial Neural Networks (AANs) atau Teknik optimasi non-linear untuk mengkombinasi atribut-atribut menjadi perkiraan dari parameter yang diinginkan.

Analisis multiatribut pada penelitian ini menggunakan kategori yang kedua. Algoritma dari multiatribut akan mencari sebuah fungsi yang akan mengkonversi *m* atribut yang berada ke dalam properti yang diinginkan, ini dapat ditulis sebagai berikut:

$$P(x, y, z) = F\,[A_1(x, y, z), \ldots\ldots, A_n(x, y, z)] \tag{2.7}$$

Dimana:

P = properti log, sebagai fungsi dari koordinat x,y,z

F = fungsi yang menyatakan hubungan antara atribut seismik dan properti log

Ai = atribut n, dimana i = 1….,n.

Untuk kasus yang paling sederhana, hubungan antara log properti dan atribut seismik dapat ditunjukkan oleh persamaan jumlah pembobotan linear.

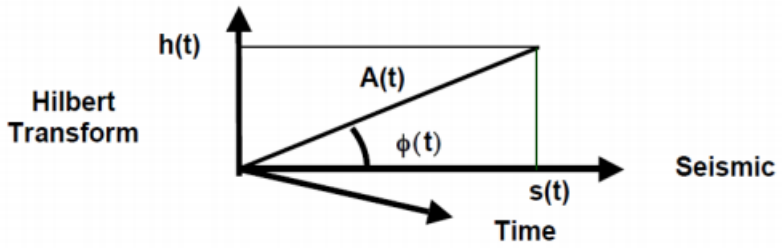
$$P = w_0 + w_1 A_1 + \ldots\ldots + w_n A_n \tag{2.8}$$

Dimana;

w_1 = nilai bobot dari n+1, dimana 1 = 0 ,…..,n

2.5.1 Atribut Sesaat (*Instantaneous Attribute*)/Atribut Kompleks

Atribut kompleks dapat diturunkan dari atribut yang merupakan atribut utama dapat ditunjukkan pada persamaan yang didapatkan dari transformasi Hilbert. Atribut ini dihitung dari trace kompleks C(t), yang terdiri dari trace seismik s(t), dan transformasi Hilbert h(t), yang merupakan pergeseran fasa sebesar 90˚ dari trace seismik.



Gambar 2. 11 Penulisan trace kompleks dalam bentuk polar (Russel, et.Al., 1997)

Penulisan tras kompleks dalam bentuk polar, seperti yang terlihat pada Gambar 2.11, memberikan kita dua atribut dasar, yaitu kuat refleksi A(t) dan fasa sesaat ϕ(t).

$$C(t) = s(t) + ih(t) \tag{2.9}$$

$$= A(t)e^{\mathrm{i}\phi(t)} \tag{2.10}$$

$$= A(t)cos\phi(t) + iA(t)sin\phi(t); i = \sqrt{-1} \tag{2.11}$$

Sehingga,

$$A(t) = \sqrt{s(t)^2 + h(t)^2} \tag{2.12}$$

$$\phi(t) = \tan^{-1}\left[\frac{h(t)}{s(t)}\right] \tag{2.13}$$

Atribut dasar yang ketiga adalah frekuensi sesaat yang merupakan turunan atau perubahan fasa sesaat terhadap waktu. Persamaan frekuensi sesaat ini dapat ditulis sebagai:

$$\omega(t) = \frac{d\phi(t)}{t} \tag{2.14}$$

Atribut lainnya yang merupakan kombinasi dari tiga atribut dasar tersebut antara lain:

$cos\phi(t)$ = kosinus fasa sesaat

$A(t)cos\phi(t)$ = *amplitude weighted cos phase*

$A(t)\phi(t)$ = *amplitude weighted phase*

$A(t)\omega(t)$ = *amplitude weighted frequency*

2.5.2 Atribut Jendela frekuensi (*windowed frequency attributes*)

Atribut ini menganalisa menggunakan window dengan proses fourier transform dari setiap tras seismik yang diambil. Window ini baik amplitudo frekuensi rata-rata maupun amplitudo frekuensi dominan digunakan dan nilainya ditempatkan ditengah-tengah window. Dalam proses nya atribut jendela frekuensi ini dipilih 32 sampel, kemudian atribut frekuensi yang baru dihitung dan demikian seterusnya.

2.5.3 Atribut *Filter Slice*

Set ketiga dari atribut terdiri dari *narrow band filter slices* dari tras seismik. Enam slice yang digunakan adalah sebagai berikut:

5/10-15/20 Hz

15/20-25/30 Hz

25/30-35/40 Hz

35/40-45/50 Hz

45/50-55/60 Hz

55/60-65/70 Hz

2.5.4 *Derivative Attributes*

Set keempat dari atribut didasarkan pada turunan pertama atau kedua dari tras seismik ataupun amplitudo sesaat. Turunan tersebut dihitung melalui cara berikut:

$$d1_i = \frac{S_i - S_{i-1}}{\Delta t} \tag{2.15}$$

$$d2_i = \frac{d1_i - d1_{i-1}}{\Delta t} = \frac{s_i - 2S_{i-1} - S_{i-2}}{\Delta t^2} \tag{2.16}$$

Dimana,

S_i= sampel tras seismik atau amplitudo sesaat ke-i

$d1_i$= turunan pertama dari i

$d2_i$= turunan kedua dari i

dt = interval sampling

2.5.5 *Integrated Attributes*

Jenis kelima dari atribut yang terdapat pada Emerge didasarkan pada integrasi dari tras seismik atau kuat refleksi. Nilai integrasi dihitung dengan cara berikut:

$$I_s = S_i + I_{i-1} \tag{2.17}$$

S_i = sampel ke-I dari tras seismik atau nilai kuat refleksinya, I_i = nilai integrasi.

2.5.6 Atribut Waktu (*Time Atribut*)

Atribut waktu ini secara sederhana merupakan nilai waktu dari tras seismik yang membentuk sebuah fungsi "ramp" yang dapat menambah sebuah trend dalam menghitung parameter reservoir, sehingga dicari hubungan antara properti log dan kombinasi atribut dari tras seismik pada interval target.

2.5.7 Cross Plotting

Salah satu cara untuk mengukur korelasi antara data target dan atribut adalah dengan melakukan cross plot antara keduanya.

Hubungan linier antara log target dan atribut ditunjukkan oleh sebuah garis lurus yang memenuhi persamaan:

$$y = a + b * x \tag{2.18}$$

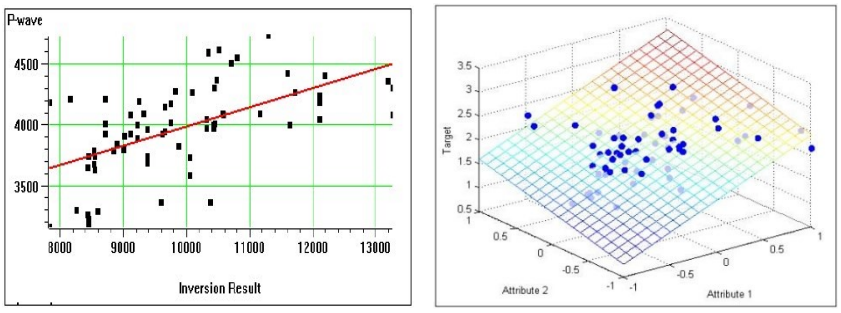
koefisien a dan b pada persaman ini diperoleh dengan meminimalisasikan *mean-square prediction error* :

$$E^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - m_x)(y_i - m_y) \tag{2.19}$$

Dimana penjumlahan dilakukan pada setiap titik di crossplot. Dengan mengaplikasikan garis regresi tersebut kita dapat memberikan prediksi untuk atribut target.

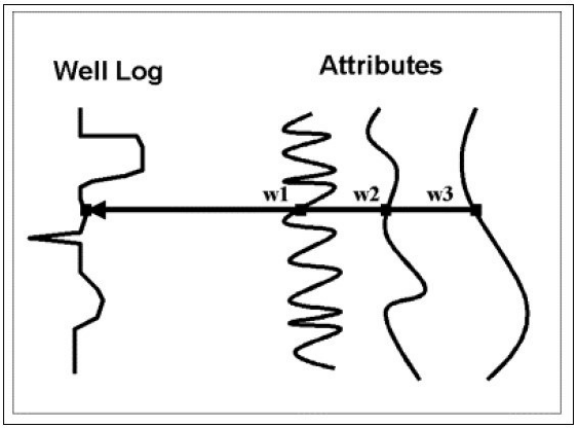
2.5.8 Regresi Linear Multiatribut

Regresi linear multiatribut merupakan pengembangan dari *cross plot* konvensional. Dalam metoda ini, tujuan kita adalah untuk mencari sebuah operator yang dapat memprediksi log sumur dari data seismik didekatnya. Pada kenyataannya, kita menganalisa data atribut seismik dan bukan data seismik itu sendiri. Salah satu alasan mengapa kita melakukan hal ini karena menggunakan data atribut seismik lebih menguntungkan daripada data seismik itu sendiri. Banyak dari atribut ini bersifat non-linear, sehingga mampu meningkatkan kemampuan prediksi.



Gambar 2. 12 Ilustrasi Cross-plot dengan menggunakan (a) satu atribut dan (b) dua atribut (Hampson & Russel, 2008)

Pengembangan (*extension*) analisis linier konvensional terhadap *multiple* atribut (regresi linier multiatribut) dilakukan secara langsung. Sebagai penyederhanaan, kita mempunyai tiga atribut seperti yang terlihat pada gambar 2.12.



Gambar 2. 13 Contoh kasus tiga atribut seismik, tiap sampel log target dimodelkan sebagai kombinasi linier dari sampel atribut pada interval waktu yang sama. (Russel, 2008).

Pada tiap sampel waktu, log target dimodelkan oleh persamaan linear:

$$L(t) = w_0 + w_1 A_1(t) + w_2 A_2(t) + w_3 A_3(t) \qquad (2.20)$$

Pembobotan (*weights*) pada persamaan ini dihasilkan dengan meminimalkan *mean-squared prediction error.*

$$E^2 = \frac{1}{N}\sum_{i=j}^{N}(L_i - w_0 - w_1 A_{1i} - w_2 A_{2i} - w_3 A_{3i})^2 \qquad (2.21)$$

Solusi untuk empat pembobotan menghasilkan persamaan normal standar:

$$\begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} N & \sum A_{1i} & \sum A_{1i} & \sum A_{1i} \\ \sum A_{1i} & \sum A^2{}_{1i} & \sum A^2{}_{1i} A_{2i} & \sum A_{1i} A_{3i} \\ \sum A_{2i} & \sum A^2{}_{1i} A_{2i} & \sum A^2{}_{2i} A_{2i} & \sum A_{2i} A_{2i} \\ \sum A_{3i} & \sum A^2{}_{1i} A_{3i} & \sum A_{2i} A_{3i} & \sum A^2{}_{3i} \end{bmatrix}^{-1} \begin{bmatrix} \sum A_{1i} \\ \sum A_{1i} L_i \\ \sum A_{2i} L_i \\ \sum A_{3i} L_i \end{bmatrix} \qquad (2.22)$$

Seperti pada kasus atribut tunggal, *mean-square error* yang dihitung menggunakan pembobotan, merupakan pengukuran kesesuaian untuk transformasi tersebut, dimana sekarang koordinat x merupakan nilai log yang diprediksi dan koordinat y merupakan nilai real dari data log.

Dalam memilih kombinasi atribut yang paling baik adalah dengan cara yang dinamakan *step-wise regression:*

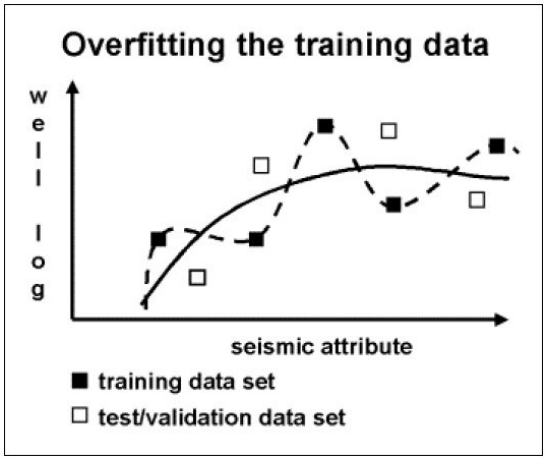
1. Dicari atribut tunggal pertama yang paling baik dengan menggunakan *trial and error*. Untuk setiap atribut yang terdapat pada *software* dihitung eror prediksinya. Atribut terbaik adalah atribut yang memberikan eror prediksi terendah. Atribut ini selanjutnya akan disebut atribut -a.
2. Dicari pasangan atribut yang paling baik dengan mengasumsikan anggota pasangan yang pertama adalah atribut-a. Pasangan yang paling baik adalah pasangan yang memberikan eror paling kecil. Atribut ini selanjutnya akan disebut atribut-b
3. Dicari tiga buah atribut yang berpasangan paling baik, dengan mengasumsikan dua buah anggota yang pertama atribut-a dan atribut-b. Tiga buah atribut yang paling baik adalah yang memberikan eror prediksi yang paling kecil. Prediksi ini berlangsung terus sebanyak yang diinginkan.

   Eror prediksi, En, untuk n atribut selalu lebih kecil atau sama dengan En-1 untuk n-1 atribut, tidak peduli atribut mana yang digunakan.

2.5.9 Validasi

Pertanyaan selanjutnya yang harus dijawab adalah kapan kita harus berhenti menambahkan jumlah atribut. Bagaimanapun juga transformasi multi-atribut dengan jumlah atribut N+1 selalu mempunya prediksi eror lebih kecil atau sama dengan transformasi dengan N atribut. Menambah jumlah atribut, serupa dengan mencari kurva regresi yang cocok untuk sebuah plot data, dengan menggunakan polinomial yang ordenya semakin besar.

Apabila jumlah atribut yang digunakan semakin banyak, maka eror prediksi akan semakin berkurang. Masalah yang kemudian timbul adalah, biasanya data yang dihasilkan akan buruk bila diterapkan pada data baru (yang tidak termasuk dalam data training), karena atribut tersebut terlalu dicocokan dengan data training. Hal ini biasa disebut dengan over-training.



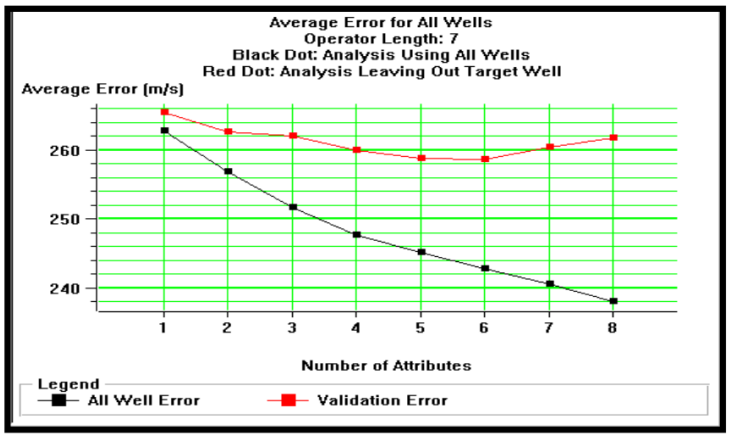
Gambar 2. 14 Ilustrasi cross-validasi. Kedua kurva digunakan untuk mencocokkan data. Kurva putus-putus menunjukkan korelasi yang baik dengan data training. Namun buruk jika kemudian digunakan set data validasi (Russel, 2008).

Untuk mengukur validasi dari jumlah atrbut yang digunakan, dilakukan prosedur sebagai berikut. Misalnya, terdapat tiga buah atribut dan lima buah well. Untuk perhitungan pertama, sumur pertama tidak diikutkan. Lalu dihitung koefisien yang dihasilkan, lalu dihitung prediksi eror dari satu sumur dengan rumus:

$$E^2 = \frac{1}{N} \sum_{i=1}^{N} (\varphi_i - w_0 - w_1 * I_i - w_2 * E_i - w_3 * F_i)^2 \quad (2.23)$$

Ini adalah eror validasi untuk sumur satu. Proses ini kemudian diulang untuk sumur 2, sumur 3 dan sumur selanjutnya. Dengan eror rata-rata dihitung:

$$E_A = \frac{(E_1+E_2+E_3+E_4+E_5)}{5} \quad (2.24)$$



Gambar 2. 15 Plot validasi total eror dan training eror (modifikasi dari Hampson et al.,2001)

gambar 2.15 diatas menunjukkan plot yang telah ditambahkan validasi total eror. Seperti yang diharapkan validasi eror untuk setiap jumlah atribut selalu lebih besar dari training eror. Hali ini disebabkan oleh penghilangan sebuah sumur dari set training selalu menurunkan hasil kemampuan prediksi. Perlu dicatat bahwa setelah atribut kedua, atribut lainnya menyumbang peningkatan kecil pada validasi eror, dan pada kenyataannya, secara bertahap menyebabkan peningkatan nilai prediksi eror. Kurva validasi eror tidak menurun secara monoton tetapi menunjukkan minimum lokal disekitar empat atribut, kemudian secara bertahap meningkat serta sistem akan mengalami over training. Pada umumnya, jika kurva validasi eror secara jelas menunjukkan paling minimum, jumlah atribut pada titik tersebut adalah optimum.